# Employing BAC-reporter constructs in the sea anemone *Nematostella vectensis*

Antje H.L. Fischer*, Sarah Tulin*, David Fredman† Joel Smith*°

*Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA; †Department of Molecular Evolution and Development, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.

°corresponding author: jcsmith@mbl.edu, Marine Biological Laboratory, 7 MBL Street, Woods Hole 02543, MA, USA, Tel: (508)-289-7253, 508-289-7900

From the symposium ‘‘Understanding First Order Phenotypes: Transcriptomics for Emerging Models’’ presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2013, in San Francisco, California.

## Abstract

Changes in the expression and function of genes drive evolutionary change. Comparing how genes are regulated in different species is therefore becoming an important part of evo-devo studies. A key tool for investigating the regulation of genes is represented by Bacterial Artificial Chromosomes-reporter constructs (BAC). BACs are large insert libraries, often >>100 kb, which thus capture the genomic sequences surrounding a gene of interest, including all, or nearly all, of the elements underpinning regulation. Recombinant BACs, containing a reporter gene in place of the endogenous coding sequence of genes, can be utilized to drive the expression of reporter genes under the regulatory control of the gene of interest while still embedded within its genomic context. Systematic deletions within the BAC reporter construct can be used to identify the minimal reporter in an unbiased way, avoiding the risk of overlooking regulatory elements that may be many kilobases away from the transcription start-site. *Nematostella vectensis* (Edwardsiidae, Anthozoa, Cnidaria) has become an important model in regenerative biology, ecology, and especially in studies of evo-devo and gene-regulatory networks due to its interesting

phylogenetic position and amenability to molecular techniques. The increasing interest in this rising model system also led to a demand for methods that can be used to study the regulation of genes in *Nematostella.* Here we present our progress in employing BAC reporter constructs to visualize gene-expression in *Nematostella.* Using a new *Nematostella*-specific recombination cassette, we made nine different BAC reporter constructs. Although five BAC recombinants gave variable effects, three constructs, namely *Nv-bra*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC, and *Nv-grm*:eGFP::L10 BAC, delivered promising results. We show that these three constructs express the reporter gene eGFP in 10.4% – 17.2% of all analyzed larvae, out of which 26.2 – 41.9% express GFP in a mosaic fashion within the expected domain. In addition to the expression within the known domains, we also observed cases of misexpression of eGFP and examples that could represent actual expression outside the described domain. Furthermore, we deep-sequenced and assembled five different BACs containing *Nv-chordin, Nv-foxa, Nv-dpp, Nv-wnta,* and *Nv-wnt1*, to improve assembly around these genes. The use of BAC reporter constructs will foster *cis*-regulatory analyses in *Nematostella* and thus help to improve our understanding of the regulatory network in this cnidarian system. Ultimately, this will advance the comparison of gene-regulation across species and lead to a much better understanding of evolutionary changes and novelties.

## Introduction

*Nematostella vectensis* is a small sea anemone that lives in estuarine habitats in North America and parts of the UK (Gilmore et al. 2013). Sea anemones are anthozoans and thus belong to the taxon Cnidaria, which includes jellyfish, box jellies, corals, and sea pens (Collins et al. 2006). Phylogenetic studies show that Cnidarians are the sister group to all bilaterians (e.g. (Collins 1998; Medina et al. 2001; Dunn et al. 2008; Hejnol et al. 2009) and therefore have an

exceptionally interesting phylogenetic position for evolutionary studies regarding the origin of many bilaterian traits (Steele et al. 2011). Recently, *Nematostella* has become an interesting model organism in studies of regenerative biology, ecology, and regulation of genes and continues to be valuable for studies of evo-devo (Hand 1995; Scholz and Technau 2003; Martindale 2004; Darling et al. 2005; Kusserow et al. 2005; Matus et al. 2006; Rentzsch et al. 2006; Matus et al. 2007a; Matus et al. 2007b; Kumburegama et al. 2008; Matus et al. 2008; Reitzel et al. 2008; Rentzsch et al. 2008; Burton and Finnerty 2009; Saina et al. 2009; Reitzel et al. 2010; Kumburegama et al. 2011; Passamaneck and Martindale 2012; Röttinger et al. 2012). In addition to its interesting phylogenetic position, the establishment of several techniques of molecular biology for *Nematostella* has made it a particularly suitable model organism (Genikhovich and Technau 2009c; Genikhovich and Technau 2009a; Genikhovich and Technau 2009b; Genikhovich and Technau 2009e; Layden et al. 2013; Stefanik et al. 2013a; Stefanik et al. 2013b; Wolenski et al. 2013). *Nematostella* is easy to keep and proliferate in the laboratory, and spawning can be induced in males and females on a regular basis (Genikhovich and Technau 2009e; Genikhovich and Technau 2009d; Stefanik et al. 2013a). The genome is sequenced and annotated and the transcriptome is available (Ryan et al. 2006; Putnam et al. 2007; Helm et al. 2013; Tulin et al. 2013). Gene function can be perturbed with morpholino oligonucleotides, overexpression of RNA, and the use of double-stranded RNA (e.g. Wikramanayake et al. 2003; Pankow and Bamberger 2007; Rentzsch et al. 2008; Kumburegama et al. 2011; Marlow et al. 2012). Transgenesis was recently reported for *Nematostella* (Renfer et al. 2010).

Directed studies of comparative development have elucidated several aspects of animal evolution during the past two decades (Love and Raff 2003). The comparison of developmental patterns of gene-expression and of whole genomes, as well as the identification of molecular fingerprints in

a variety of different cell types, has enhanced our understanding of evolutionary processes (Arendt 2008; Carroll 2008; Sommer 2009). At the same time, it has become clear that evolutionary novelties are frequently the result of changes in gene regulation and not solely based on newly evolving genes or changes in coding sequence (summarized by Wray 2007; Carroll 2008).

Stern (1998) published one of the first studies suggesting that differences in gene regulation cause species-specific morphological differences. The patterning of special bristles on the second leg of flies differs among *Drosophila melanogaster, D. simulans,,* and *D. virilis* and is caused by variations in the cis-regulatory region of the Hox gene *Ultrabithorax* (Stern 1998). Another example is the convergent evolution in pigmentation of the wings in *Drosophila* species, which is caused by independent mutations within the same cis-regulatory element (2006).

Interesting examples of evolutionary changes in cis-regulatory elements are also observed in vertebrates. By crossing three-spine stickle back species that had dorsal spines and a pelvic girdle with species in which these structures were reduced, Shapiro et al. (2004) mapped the genetic basis for pelvic reduction to differential regulation of *paired-like homeodomain transcription factor 1* (*pitx1*). These examples illustrate how the understanding of gene regulation enhances our understanding of evolutionary changes; the comparison of gene-regulatory networks across species is becoming an important part of evo-devo studies (Fischer and Smith 2012).

To investigate gene regulation and cis-regulatory elements, Bacterial Artificial Chromosomes-reporter constructs (BAC) are a key molecular tool. BACs contain large insert libraries capturing the genomic context of a gene of interest (GOI) that can be engineered to drive the expression of reporter genes under the regulatory control of that gene (Smith 2008; Smith and Davidson 2008a;

Smith and Davidson 2008b; Smith and Davidson 2009; Wahl et al. 2009; de-Leon and Davidson 2010; Damle and Davidson 2011; Ohguro et al. 2011; Materna and Davidson 2012). Once the normal expression of the reporter construct has been established, a deletion series within the reporter construct is used to identify the minimal reporter, a small region, perhaps 1-2 kb in length, containing all the necessary information to drive spatially and temporally correct activity of the reporter (Smith 2008). Within this reduced search space, transcription-factor binding sites can be identified by phylogenetic footprinting and the function of individual binding sites can be tested via PCR-directed mutations (Smith 2008). Finally, developmental gene-regulatory networks and subcircuits can be compared between species to understand evolutionary changes in gene regulation. One example, in which BAC reporter constructs have been employed to foster comparison of the developmental gene-regulatory networks, is the cis-regulatory analyses of *otxβ1/2* in starfish and sea urchins. Hinman et al. (2007) and McCauley et al. (2010) showed that the regulatory input of three out of four genes controlling the expression of *otxβ1/2* is conserved between starfish and sea urchins. The fourth gene that regulates the expression of *otxβ1/2* in starfish, *tbrain*, is not conserved but is co-opted to regulate the development of the embryonic skeleton in sea urchins, which is a lineage-specific evolutionary novelty (Hinman et al. 2007).

Here we present our progress on using BAC reporter constructs in *Nematostella.* We show that they can faithfully express the reporter gene in a mosaic fashion within the expected domain. We engineered BAC reporter constructs and want to make them available and encourage their use as a powerful tool in cis-regulatory analyses. Finally, we sequenced and assembled five different BACs, to validate the genomic architecture surrounding the respective regions of these genes.

**Results:**

Our goal was to ascertain whether recombineered BAC-reporter constructs could be employed to visualize and analyze gene-expression in *Nematostella* embryos and planula larvae. Our experiments show that BAC reporters can reliably express GFP in *Nematostella* and thus provide a promising tool for *cis*-regulatory analyses, live imaging, and gene-expression studies.

Sequencing *Nematostella BACs for chordin, dpp, foxa, wntA*, and *wnt1*

We received a copy of the *Nematostella* BAC library containing about 27000 clones with an average insert size of 178 kb from http://bacpac.chori.org/library.php?id1⁄4219. Through a screen using digoxigenin-labeled DNA probes, we identified five BACs that each contained one of the following genes: *Nv-chordin, Nv-dpp, Nv-foxa, Nv-wnta*, and *Nv-wnt1*. Before constructing BAC reporters for each of these clones, we first sought to validate the genomic architecture surrounding these genes. To this end, we sequenced the above BAC clones using the Illumina HiSeq 1000 platform. We acquired 250,775,371 bp total sequences (Quality control analysis presented in Supplemental File 1). Using CLC Genomics Workbench 5.1 (CLC Bio), we first trimmed the reads using default parameters to remove adapters and any low quality sequence (see Methods), which left us with 173,107,347 bp. We then assembled the trimmed reads *de novo.* The full assembly is publically available on the WHOAS server [digital object identifier (DOI) 10.1575/1912/6068]. Finally, we mapped the *de novo* assembled contigs to the *Nematostella* genome using the CLC mapping algorithm and visualized the alignments in the Integrative Genomics Viewer 2.3. We were able to confirm that all of the selected BACs contain a substantial length of genomic sequence (>20 kb) both upstream and downstream of the transcription start site of each gene of interest (GOI).

We also identified four additional BAC clones for four more GOI using a computational approach (see methods). These BAC clones contained *Nv-brachyury, Nv-gremlin, Nv-snailB*, and *Nv-tolloid* respectively including 20 kb of genomic DNA sequence upstream and downstream. The exact position for these BACs within the genome and their estimated length are provided in Table 2.

The recombination cassette and reporter construction

In order to introduce a reporter gene into the BACs we produced a recombination cassette containing the following: the coding sequence for the fusion protein of eGFP::L10, 174 bp of the 3'UTR of *Nv-L10* including a termination signal, and a Kanamycin resistance flanked by FRT-sites (Fig. 1A).

While the eGFP protein serves as a reporter that can be visualized *in vivo*, the ribosomal protein *Nv-L10* (L10) was included for future studies in a "BACarray" approach as described by Heiman et al. (2008) and Jiao and Meyerowitz (2010). In this approach, the authors take advantage of the fusion protein between L10 and eGFP (Heiman et al. 2008; Jiao and Meyerowitz 2010). As the ribosomal protein L10 becomes part of the ribosome, the eGFP can be used as an epitope to pull down the ribosome and with it mRNAs that are being translated at the time of sampling. In a next step, the mRNAs are identified by sequencing and thus translational profiling can be carried out on all eGFP::L10-expressing cells (Heiman et al. 2008; Jiao and Meyerowitz 2010).

Because BAC clones are too large to manipulate by standard subcloning methods, the BAC reporters were constructed via homologous recombination (Heintz 2001; Lee et al. 2001; Lee et al. 2007). An overview of the procedure is provided in Figure 1A-D. We made BAC reporter

constructs for the following genes: *Nv-bra*:eGFP::L10 BAC, *Nv-chd*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC, *Nv-foxa*:eGFP::L10 BAC, *Nv-grm*:eGFP::L10 BAC, *Nv-gsc*:eGFP::L10 BAC, *Nv-snailB*:eGFP::L10 BAC, *Nv-tld*:eGFP::L10 BAC, and *Nv-wnt1*:eGFP::L10 BAC.

To prepare BAC reporters for microinjection we linearized the BAC reporter constructs with NotI, and prepared injection solutions containing approximately 250 molecules/pl of the constructs in a balanced salt solution together with the injection tracer Rhodamine Dextran and injected them into *Nematostella* zygotes. We examined the larvae at 27 – 56 hours post fertilization (hpf) and recorded how many of the injected animals showed eGFP expression. We limited our analyses to animals younger than three days, since the expression of endogenous eGFP can severely obstruct the analyses in older stages. Other fluorescent reporters may be employed if later developmental stages need to be analyzed.

The expression of BAC reporter constructs in *Nematostella* planula larvae

We did not observe any toxicity due to injecting recombinant BACs and the injected zygotes developed normally into adult animals. The *Nv-bra*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC and *Nv-grm*:eGFP::L10 BAC resulted in frequent eGFP-expression.

After injecting the *Nv-bra*:eGFP::L10 BAC a total of 11.2% of all injected larvae that were analyzed expressed eGFP (Table 3). Whole-mount *in-situ* hybridizations (WMISH) showed that *Nv-brachyury* is expressed in a ring around the blastopore in late gastrula and early planula stages (Fig. 2) (Scholz and Technau 2003). Of all larvae injected with the *Nv-bra*:eGFP::L10 BAC, 41.9% showed eGFP expression in domains of varying sizes around the blastopore (Table 3; Fig. 2). These domains contained reporter gene-expression and thus represented clones within the

expression pattern as described from WMISH. In some larvae more than one small eGFP-expression domain could be observed near the blastopore (Fig. 2D). We observed eGFP-expression outside the region of the blastopore in 58.1% of all injected larvae, indicating potential overexpression of the construct (Table 3).

*Nv-dpp* mRNA has a complex dynamic pattern of expression throughout embryonic and larval development. During early gastrula stages, *Nv-dpp* mRNA is expressed around the blastopore, and then becomes restricted to a small domain neighboring the blastopore (Matus et al. 2006; Rentzsch et al. 2006; Saina et al. 2009). At later gastrula and planula stages, *Nv-dpp* expression appears in a rough stripe along the oral-aboral axis in the endoderm, and ends in a spot-like domain neighboring the blastopore (Fig. 3) (Matus et al. 2006; Rentzsch et al. 2006; Saina et al. 2009). After injecting the *Nv-dpp*:eGFP::L10 BAC, we noticed eGFP expression in 10.4% of all analyzed larvae (Table 3). The eGFP expression recapitulated in a mosaic fashion, the expression domains identified from WMISH in 31.3% of all GFP-positive larvae (Table 3). In two-day-old planula larvae, we observed eGFP expression along an oral-aboral stripe in several small domains (Fig. 3). However, while WMISH showed the asymmetric strip-like expression only within the endoderm during normal development, we observed eGFP expression also within the ectoderm, in domains variable size, in 68.8% of all eGFP-expressing larvae (Fig. 3; Table 3). Currently, we are not sure if this represents additional expression domains that have not yet been reported or an artifact of overexpression of the BAC reporter construct.

After injecting the *Nv-grm*:eGFP::L10 BAC, 17.2% of all larvae that were analyzed showed eGFP expression (Table 3). Thus, despite the fact that the *Nv-grm*:eGFP::L10 BAC was injected at a similar concentration as the *Nv-bra*:eGFP::L10 BAC and the *Nv-dpp*:eGFP::L10 BAC, the

frequency of eGFP expression was approximately 1.5-1.8 fold higher for the *Nv-grm* BAC than for the others. Based on WMISH, *Nv-grm* mRNA is first expressed in the aboral endoderm on one side of the embryo, shortly after gastrulation (Rentzsch et al. 2006). During the mid-planula stage *Nv-gremlin* mRNA forms a gradient within the endoderm along an aboral to oral axis (Fig. 4) (Rentzsch et al. 2006). The eGFP expression we observed in 26.2% of the injected larvae partially recapitulated the published expression patterns, but appeared to be highly variable (Table 3). In 73.8% of all eGFP-positive larvae, we saw eGFP expression in the ectoderm, outside the expression domain indicated by WMISH (Fig. 4; Table 3).

Among the other recombineered BAC reporters we have made, we observed eGFP expression for:

- *Nv-chd*:eGFP::L10 BAC
- *Nv-foxa*:eGFP::L10 BAC
- *Nv-gsc*:eGFP::L10 BAC
- *Nv-snailB*:eGFP::L10 BAC, and
- *Nv-wnt1*:eGFP::L10 BAC

Although eGFP expression after the injection of these constructs was observed in preliminary experiments, these results have yet to be repeated. While we extol the benefits of working with BAC reporters, there are numerous technical challenges, not least being the preparation and storage of very long, linearized DNA. We are currently re-preparing these reporters for further validation. By contrast, we have never observed any eGFP expression after injecting the *Nv-tld*:eGFP::L10 BAC, which may indicate a problem with the construction.

## Discussion

To our knowledge, this report represents the first time the use of large reporter constructs has been documented in *Nematostella*. We found that the expression of the reporter gene eGFP that was introduced via homologous recombination into BACs varied, depending on the gene or BAC clone respectively. For three constructs, namely *Nv-bra*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC, and *Nv-grm*:eGFP::L10 BAC, we obtained promising results that encourage us to pursue this approach. However, the amount of ectopic expression of eGFP as shown above, the low abundance or complete lack of expression for several other reporter constructs shows that there are problems remaining that need to be addressed.

BAC reporter constructs have proven to be powerful tools for visualizing gene-expression in live embryos and larvae including their use for cis-regulatory analyses (e.g. (Jeong et al. 2006; Gebhard et al. 2007; Smith and Davidson 2008a; Smith and Davidson 2008b; Smith and Davidson 2009; Damle and Davidson 2011). The GOI remains within its genomic landscape, which includes all its enhancer elements and repressors even those that are many kilobases away. It is therefore highly probable that all the regulatory information is present within the construct (Smith 2008). *Cis*-regulatory elements can be identified by systematic deletion experiments within the BAC reporter construct (Smith 2008). Furthermore, BAC reporter constructs can be used in live-imaging approaches to visualize gene-expression *in vivo* (Smith 2008; Smith and Davidson 2008a; Smith and Davidson 2008b; Smith and Davidson 2009; Wahl et al. 2009; de-Leon and Davidson 2010; Damle and Davidson 2011; Ohguro et al. 2011; Materna and Davidson 2012), or to drive RNAi expression in specific cell types (Smith and Davidson 2008c). Finally, the use of eGFP::L10 fusion proteins in BACarray assays as described by Heiman et al. (2008) and Jiao and Meyerowitz (2010) provides a powerful tool for translation profiling.

The wide range of applications of BAC reporter constructs encouraged us to explore their use as a tool for investigating gene regulation and for visualizing gene-expression in the sea anemone *Nematostella vectensis*. As with any powerful technique, there are a number of considerations. We focus on the key ones below.

Mosaic expression of the reporter construct

In several cases the expression of eGFP after injecting *Nv-bra*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC, or *Nv-grm*:eGFP::L10 BAC is faithful within the expression domains described by WMISH (Scholz and Technau 2003; Matus et al. 2006; Rentzsch et al. 2006). However, although the eGFP domains were faithful in some cases, they were mostly mosaic and recapitulated only a subsection of the endogenous gene-expression domain.

In injected embryos, large DNA fragments are stably integrated into one or more of the blastomeres during early cleavage (de-Leon and Davidson 2010). This integration results in a random mosaic incorporation pattern and thus a clone of cells shows the expression of the reporter gene within the appropriate domain (de-Leon and Davidson 2010). Most likely, this is also the case in *Nematostella*, resulting in a mosaic eGFP expression. To understand this in more detail, a systematic analysis of several injected embryos and larvae will be necessary to obtain a comprehensive image of the expression of the reporter gene.

Accuracy and sensitivity of reporter constructs

In addition to the mosaic expression within the known gene-expression domains, we also observed apparent misexpression and overexpression of eGFP outside the described gene-

expression domain. Misexpression and overexpression of reporter constructs can have several causes. Injecting too much of a BAC construct is a potential cause and can be addressed by testing a range of concentrations. An abnormally high rate of injected larvae expressing the transgene would be an indication of over injection. We obtained almost twice as many eGFP-expressing larvae after the injection of the *Nv-grm*:eGFP::L10 BAC in comparison with the *Nv-bra*:eGFP::L10 and the *Nv-dpp*:eGFP::L10 BACs although all three constructs were injected at similar concentrations. We also observed more larvae that showed eGFP expression outside the published *gremlin* expression domain (73.8%) as determined by WMISH, suggesting overexpression of the *Nv-grm*:eGFP::L10 BAC. Lowering the concentration of *Nv-grm*:eGFP::L10 BAC might also lower the percentage of eGFP-expressing larvae and the amount of misexpression of the construct. Therefore, we suggest that a suitable concentration should be identified individually for each construct.

Nevertheless, even given the above concerns with overexpression, if a gene is expressed in some domains at high levels and in others at low levels during the same developmental stage, domains of weak gene-expression might not always be picked up by WMISH and thus could be overlooked. The sensitivity of fluorescent reporters may be important in this respect. One example is the expression of *HesC* in sea urchins. When first described, *hesC* was reported to be expressed everywhere except in the skeletogenic mesoderm, specifically the *delta*-expressing cells in the center of the vegetal plate at 12 hpf in unhatched blastulae (Revilla-i-Domingo et al. 2007). Using *Sp-hesC*:GFP BAC reporter constructs, Smith and Davidson (2008b) gave a more detailed description of *hesC* expression in the sea urchin, subdividing it into two domains with low-level expression, and one domain with a high-level expression in the mesenchyme blastula stage.

Endogenous expression of fluorescent proteins in *Nematostella*

Nowadays a great variety of fluorescent proteins are available (Shaner et al. 2005). We chose eGFP as a fluorescent reporter gene since it was successfully used in numerous studies of gene-expression in sea urchins and appeared to have suitable kinetic properties such as its maturation rate and turnover rate (Smith 2008).

Low levels of green and red fluorescence can occur naturally in *Nematostella* embryos and planula larvae (Supplemental File S2) (Ikmi and Gibson 2010). However, this low-level endogenous expression does not hamper the analyses of distinct expression of GFP. Previous studies have shown that GFP can be used as a reporter gene in *Nematostella.* Wikramanayake et al. (2003) successfully injected *Nv-ß-catenin*:GFP mRNA into zygotes to visualize the nuclearization of ß-catenin and the effect of LiCl treatments *in vivo* in *Nematostella* embryos. Strong green auto-fluorescence occurs in the pharyngeal region in primary polyps (Ikmi and Gibson 2010), restricting the use of green fluorescent proteins to younger animals. In addition to green auto-fluorescence, the endogenous fluorescent protein, *Nvfp7*, is also expressed in parts of the endoderm at later stages (Ikmi and Gibson 2010).

Potential integration of germ lines

Depending on the experimental design, stable transgenic lines can be advantageous over the transient expression of reporter constructs, especially if later developmental stages are of interest or if selected cell populations are to be traced through several stages (Renfer et al. 2010). Using the I-SceI meganuclease system, Renfer et al. (2010) generated the first stable transgenic line in *Nematostella*.

BAC reporter constructs are used as a tool to create transgenic mice and recapitulate gene-

expression patterns while avoiding positional effects of integration (e.g.Heintz 2001; Rowan and Cepko 2004; Gebhard et al. 2007). Using BAC reporter constructs takes advantage of the fact that large linearized DNA constructs such as BACs get integrated into the genome of one or more blastomeres after injecting them into zygotes (de-Leon and Davidson 2010). In contrast to the cnidarian *Hydra,* circular or linear plasmid DNA integrates only inefficiently in *Nematostella* (Renfer et al. 2010). However, DNA molecules as large as BACs might show a different frequency of integration in early *Nematostella* blastomeres and thus it remains open whether or not they will lead to stable transgenesis.

Alternative strategies

The activity of eGFP after injecting *Nv-bra*:eGFP::L10 BAC, *Nv-dpp*:eGFP::L10 BAC, and *Nv-grm*:eGFP::L10 BAC shows that BAC reporter constructs can be used in *Nematostella*. However, the mosaic expression of reporters, the in some cases large fraction of animals that show potential overexpression of the BAC reporter construct, and the variable expression of others (such as observed with *Nv-chd*-:eGFP::L10 BAC, *Nv-foxa*:eGFP::L10 BAC, *Nv-gsc*:eGFP::L10 BAC, *Nv-snailB*:eGFP::L10 BAC, *Nv-tld*:eGFP::L10 BAC, and *Nv-wnt1*:eGFP::L10 BAC) show that the use of BAC reporter constructs remains challenging in *Nematostella*. Consequently, alternative strategies should be considered for a given experiment, as they may be faster or of superior read-out in comparison to BAC reporters. In some cases, it might be beneficial to systematically test 20 to 30 kb of genomic DNA upstream and downstream of any given gene in five to ten kb segments in a reporter gene assay employing minimal promoters. The I-SceI meganuclease system as described by Renfer et al. (2010) can be used to create stable transgenic lines for several reporter constructs in laboratories that have sufficiently large facilities to host several *Nematostella* strains.

## Methods

Cloning of the recombination cassette

PCR was used to amplify *eGFP*, Kanamycin-resistance *knR* flanked by FRT-sites from pGFK (plasmid containing eGFP and Kanamycin-resistance flanked by FLT-sites, sequence provided in supplemental file S3) and *Nv-L10* with 174bp of 3' UTR from cDNA prepared from a variety of developmental stages of *Nematostella vectensis* using the SMARTer cDNA PCR synthesis kit (Clonetech, Laboratories, Inc., Mountain View, CA, USA). All primer sequences are given in supplemental file S4. Fragments were purified using the PCR purification kit (Qiagen, Valencia, CA, USA). Fusion PCR using the Phusion HighFidelity PCR master mix (New England BioLabs, Inc., Ipswich, MA, USA) resulted in an in-frame insertion of eGFP and *Nv-L10*. The eGFP::*L10* fusion fragment was amplified and fused with the FRT-site-flanked *knR* fragment via fusion PCR. The resulting recombination cassette was subcloned into PCRII Topo (Life Technology, Grand Island, NY, USA), amplified, and then cloned into pGEM (Promega, Fitchburg, WI, USA). A detailed description of the synthesis of the construct is available upon request. The sequence of the recombination cassette is provided in supplemental file S5.

Identification and homologous recombination of bacterial artificial chromosomes (BACs).

The *Nematostella* BAC libraries (http://bacpac.chori.org/library.php?id1⁄4219) were screened using digoxigenin-labeled DNA probes as described by Lee et al. (2007) to identify BACs containing the following genes: *Nv-chordin, Nv-dpp, Nv-foxa, Nv-wnta,* and *Nv-wnt1*. The sequence of the primers to make the probes is available in supplemental file S4.

To identify BACs containing *Nv-brachyury, Nv-gremlin, Nv-snailB*, or *Nv-tolloid* we BLASTed the BACend sequences (kindly provided by Prof. U. Technau, University of Vienna, Austria and

Reinhard Richard, Max-Planck Genome centre Cologne) against the *Nematostella* genome. The BAC-end sequences were trimmed of the vector sequence by BLAST comparison to UniVec, and low-quality bases with >0.05% error rate were removed. The sequences were mapped to the *Nematostella vectensis* genome (http://genome.jgi-psf.org/Nemve1/Nemve1.home.html) (Putnam et al. 2007) by BLAT using soft masking of repeats and allowing for large gaps. We retained hits that were within 2% score of the best hit, had a minimal coverage of 40% and minimum alignment identity of 85%, to allow ambiguous base calls that by BLAT are counted as mismatches. BAC end read pairs were used to identify BACs overlapping GOI.

Homologous recombination with the recombination cassette was performed as described by Yu et al. (2000) and Lee et al. (2001, 2007). The primers that were used to generate the recombination cassette with the gene-specific overhangs are provided in supplemental file S4. The correct integration site was confirmed via PCR.

The purified re-engineered BACs were linearized with NotI (New England BioLabs, Inc., Ipswich, MA, USA) and purified via drop dialyses (Smith 2008). Co-injection of 250 molecules/pl linearized BAC reporter construct and rhodamine dextran was carried out in a final concentration of 120 mM KCl. Injected embryos were cultured in 50% seawater at room temperature for one to three days.

Preparation of the sequencing libraries, Illumina-Sequencing, and Bioinformatics

Minipreps were performed on BACs containing *Nv-chd*, *Nv-foxA*, *Nv-dpp*, *Nv-wntA*, and *Nv-wnt1* as described by Warming et al. (2005). Fifty ng of BAC DNA was used to make library preparations with the Nextera kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol. After purifying the libraries with a DNA cleanup and concentrator kit (Qiagen,

Valencia, CA, USA) following the manufacturer's protocol, the libraries were size-selected for 350 bp fragments on 2% PippinPrep cassettes, followed by an additional cleanup purification (Zymo, Irvine, CA, USA).

The libraries were sequenced on an Illumina HiSeq1000 platform as 100bp single end reads. The raw reads were analyzed using CLC Genomics Workbench 5.1 (CLC Bio, Aarhus, Denmark) Next Generation Sequencing toolbox. The reads were first trimmed using the default trimming parameters: quality scores with a limit of .05, maximum number of ambiguities = 2, and the Illumina adapter list. Then the trimmed reads were assembled using the *de novo* assembler in the NGS toolbox. The assembled contigs were mapped to the *Nematostella* vectensis reference genome and converted into a track file in sam format for visualization with programs such as the Broad Institute's Integrative Genomics Viewer (http://www.broadinstitute.org/igv/). The trimmed reads were also mapped to the genome and formatted into a track as a sam file. The raw un-trimmed reads, the mapping of the trimmed reads, the assembled contig sequences and the mapping of the contigs are available under DOI 10.1575/1912/6068.

Whole-mount in situ hybridization

Planula larvae that were obtained from the same spawning event as injected larvae were fixed at 2 days post fertilization, thus at the time when the eGFP-expression was examined in injected planula larvae. Fixation and whole-mount in situ hybridization were performed as described in Wolenski et al. (2013), with the exception that the larvae were only kept in fixative one for 90 seconds on ice instead of 15 min at room temperature and after denaturing the probes for 5 minutes at 90 °C, probes were put on ice immediately and not kept at room temperature for 1 minute.

The *dpp*-clone was provided by Michael J. Layden (Martindale Laboratory, Whitney Marine Laboratory for Marine Science, University of Florida). *Nv-gremlin* and *Nv-brachyury* were cloned from cDNA prepared from a variety of developmental stages of *Nematostella vectensis* using the SMARTer cDNA PCR synthesis kit (Clonetech Laboratories, Inc., Mountain View, CA, USA). All primer sequences are provided in supplemental file S4. The *Nv-gremlin* and *Nv-brachyury* probes were prepared using the Fermentas TranscriptAid* T7 High Yield Transcription Kit (Thermo Fisher, Waltham, MA, USA). The *Nv-dpp* probe was synthesized as described in Tessmar-Raible et al. (2005).

Animal husbandry and Imaging

*Nematostella* were cultured as previously described (Hand and Uhlinger 1992). Spawning was induced as described by Fritzenwanker and Technau (2002).

Living embryos and larvae were examined using the Eclipse 50i microscope (Nikon, Tokyo, Japan) and imaged with a Digital sight DS-2MBWc (Nikon) and a TV lens 0.55x DS (Nikon, Tokyo, Japan). WMISH were imaged on a Zeiss Axioscope (Oberkochen, Germany). Fluorescent images from different channels were overlaid in ImageJ and if necessary aligned using the "Align RGB planes" plug-in. Images were rotated and cropped in Photoshop CS4 (Adobe, San Jose, CA, USA). Figure plates were prepared in Illustrator CS4 (Adobe, San Jose, CA, USA).

**Funding**

The Symposium "Understanding First Order Phenotypes: Transcriptomics for Emerging Models" was supported by EDEN. The authors were supported by institutional funds from the Marine Biological Laboratory.

**Acknowledgements**

We would like to thank Molly Phillips for very valuable help with the WMISH, great animal care and help with the injections. We are particularly thankful to Dr. Michael Layden and Prof. Mark Q. Martindale (Whitney Marine Laboratory for Marine Science, University of Florida) for providing the *Nv-dpp* cDNA clone for WMISH probe preparation. Furthermore, we would like to express our gratitude to Kasia Hammer (MBL), for making the BAC reporter constructs, help with the BAC-filter screen, and for animal husbandry. We would also like to thank Julie Hahn (Caltech, California) for training and advice on homologous BAC recombination. We would like to thank Joseph Vineis, Dr. Hilary Morrison, and Richard Fox from the Bay Paul Center, MBL, for their excellent assistance with sequencing and computation. For advice on animal handling and culturing we are especially thankful to Prof. Adam Reitzel (UNC, Charlotte, North Carolina), Dr. Heather Marlow (EMBL, Heidelberg, Germany), and Prof. Casey Dunn and Rebecca Helm (Brown University, Providence, Rode Island). We are particularly grateful to Prof. U. Technau (University of Vienna) and Reinhard Richard, (Max-Planck Genome centre Cologne) for providing the BAC-end sequences.

## Authors' contribution

AHLF cloned the recombination cassette, injected, analyzed, and imaged the larvae, cloned Nv-bra and Nv-grem, synthesized all WMISH-probes and performed and imaged the WMISH. AHLF wrote the first draft of the manuscript and prepared the figures and the illustration. ST performed the BAC filter screen, the library preparations and the computational analysis of the sequencing results. DF trimmed, analyzed and mapped the BAC-end sequences against the Nematostella genome. AHLF and JS conceived the study. JS contributed substantially to the interpretation of data and to the writing of the manuscript. All authors read and approved the final manuscript.

**Literature:**


Arendt, D. (2008). "The Evolution of Cell Types in Animals: Emerging Principles from Molecular Studies." Nature Reviews Genetics 9: 868-882.

Burton, P. M. and Finnerty, J. R. (2009). "Conserved and Novel Gene Expression between Regeneration and Asexual Fission in Nematostella Vectensis." Development genes and evolution 219(2): 79-87.

Carroll, S. B. (2008). "Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution." Cell 134(1): 25-36.

Collins, A. G. (1998). "Evaluating Multiple Alternative Hypotheses for the Origin of Bilateria: An Analysis of 18s Rrna Molecular Evidence." Proceedings of the National Academy of Sciences 95(26): 15458-15463.

Collins, A. G., Schuchert, P., Marques, A. C., Jankowski, T., Medina, M. and Schierwater, B. (2006). "Medusozoan Phylogeny and Character Evolution Clarified by New Large and Small Subunit Rdna Data and an Assessment of the Utility of Phylogenetic Mixture Models." Systematic Biology 55(1): 97-115.

Damle, S. and Davidson, E. H. (2011). "Precise Cis-Regulatory Control of Spatial and Temporal Expression of the Alx-1 Gene in the Skeletogenic Lineage of *S. Purpuratus*." Developmental Biology 357(2): 505-517.

Darling, J., Reitzel, A., Burton, P., Mazza, M., Ryan, J., Sullivan, J. and Finnerty, J. (2005). "Rising Starlet: The Starlet Sea Anemone, Nematostella Vectensis." Bioessays 27(2): 211-221.

de-Leon, S. B.-T. and Davidson, E. H. (2010). "Information Processing at the Foxa Node of the Sea Urchin Endomesoderm Specification Network." Proceedings of the National Academy of Sciences 107(22): 10103-10108.

Dunn, C. W., Hejnol, A., Matus, D. Q., Pang, K., Browne, W. E., Smith, S. A., Seaver, E., Rouse, G. W., Obst, M., Edgecombe, G. D., Sorensen, M. V., Haddock, S. H. D., Schmidt-Rhaesa, A., Okusu, A., Mobjerg Kristensen, R., Wheeler, W. C., Martindale, M. Q. and Giribet, G. (2008). "Broad Phylogenomic Sampling Improves Resolution of the Animal Tree of Life." Nature 452: 745-749.

Fischer, A. H. and Smith, J. (2012). "Evo-Devo in the Era of Gene Regulatory Networks." Integrative and Comparative Biology.

Fritzenwanker, J. H. and Technau, U. (2002). "Induction of Gametogenesis in the Basal Cnidarian Nematostella Vectensis (Anthozoa)." Development genes and evolution 212(2): 99-103.

Gebhard, S., Hattori, T., Bauer, E., Bösl, M. R., Schlund, B., Pöschl, E., Adam, N., de Crombrugghe, B. and von der Mark, K. (2007). "Bac Constructs in Transgenic Reporter Mouse Lines Control Efficient and Specific Lacz Expression in Hypertrophic Chondrocytes under the Complete Col10a1 Promoter." Histochemistry and cell biology 127(2): 183-194.

Genikhovich, G. and Technau, U. (2009a). "Anti-Acetylated Tubulin Antibody Staining and Phalloidin Staining in the Starlet Sea Anemone Nematostella Vectensis." Cold Spring Harbor Protocols 2009(9): pdb.prot5283.

Genikhovich, G. and Technau, U. (2009b). "Bromodeoxyuridine Labeling of S-Phase Nuclei in the Starlet Sea Anemone Nematostella Vectensis." Cold Spring Harbor Protocols 2009(9): pdb.prot5284.

Genikhovich, G. and Technau, U. (2009c). "In Situ Hybridization of Starlet Sea Anemone (Nematostella Vectensis) Embryos, Larvae, and Polyps." Cold Spring Harbor Protocols 2009(9): pdb.prot5282-pdb.prot5282.

Genikhovich, G. and Technau, U. (2009d). "Induction of Spawning in the Starlet Sea Anemone Nematostella Vectensis, in Vitro Fertilization of Gametes, and Dejellying of Zygotes." Cold Spring Harbor Protocols 2009(9): pdb.prot5281.

Genikhovich, G. and Technau, U. (2009e). "The Starlet Sea Anemone Nematostella Vectensis: An Anthozoan Model Organism for Studies in Comparative Genomics and Functional Evolutionary Developmental Biology." Cold Spring Harbor Protocols 2009(9): pdb.emo129.

Gilmore, T., Tarrant, A. and Finnerty, J. (2013). "A Report from the Second Nematostella Vectensis Research Conference." Development genes and evolution.

Hand, C. (1995). "Asexual Reproduction by Transverse Fission and Some Anomalies in the Sea Anemone Nematostella Vectensis." Invertebrate Biology.

Hand, C. and Uhlinger, K. R. (1992). "The Culture, Sexual and Asexual Reproduction, and Growth of the Sea Anemone Nematostella Vectensis." The Biological Bulletin 182(2): 169-176.

Heiman, M., Schaefer, A., Gong, S., Peterson, J., Day, M., Ramsey, K., Suárez-Fariñas, M., Schwarz, C., Stephan, D. and Surmeier, D. (2008). "A Translational Profiling Approach for the Molecular Characterization of Cns Cell Types." Cell 135(4): 738-748.

Heintz, N. (2001). "Bac to the Future: The Use of Bac Transgenic Mice for Neuroscience Research." Nature Reviews Neuroscience 2(12): 861-870.

Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., Martinez, P., Baguñà, J., Bailly, X. and Jondelius, U. (2009). "Assessing the Root of Bilaterian Animals with Scalable Phylogenomic Methods." Proceedings of the Royal Society B: Biological Sciences 276(1677): 4261-4270.

Helm, R. R., Siebert, S., Tulin, S., Smith, J. and Dunn, C. W. (2013). "Characterization of Differential Transcript Abundance through Time During Nematostella Vectensis Development." BMC Genomics 14(1): 266.

Hinman, V. F., Nguyen, A. and Davidson, E. H. (2007). "Caught in the Evolutionary Act: Precise Cis-Regulatory Basis of Difference in the Organization of Gene Networks of Sea Stars and Sea Urchins." Developmental Biology 312(2): 584.

Ikmi, A. and Gibson, M. C. (2010). "Identification and in Vivo Characterization of Nvfp-7r, a Developmentally Regulated Red Fluorescent Protein of Nematostella Vectensis." PLoS ONE 5(7): e11807.

Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M. and Epstein, D. J. (2006). "A Functional Screen for Sonic Hedgehog Regulatory Elements across a 1 Mb Interval Identifies Long-Range Ventral Forebrain Enhancers." Development 133(4): 761-772.

Jiao, Y. and Meyerowitz, E. M. (2010). "Cell-Type Specific Analysis of Translating Rnas in Developing Flowers Reveals New Levels of Control." Molecular systems biology 6(1).

Kumburegama, S., Wijesena, N. and Wikramanayake, A. H. (2008). "Detecting Expression Patterns of Wnt Pathway Components in Nematostella Vectensis Embryos." Methods in molecular biology (Clifton, NJ) 469: 55-67.

Kumburegama, S., Wijesena, N., Xu, R. and Wikramanayake, A. H. (2011). "Strabismus-Mediated Primary Archenteron Invagination Is Uncoupled from Wnt/B-Catenin-Dependent Endoderm Cell Fate Specification in Nematostella Vectensis (Anthozoa, Cnidaria): Implications for the Evolution of Gastrulation." EvoDevo 2(1): 2.

Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H. A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M. Q. and Holstein, T. W. (2005). "Unexpected Complexity of the Wnt Gene Family in a Sea Anemone." Nature 433(7022): 156-160.

Layden, M. J., R√∂ttinger, E., Wolenski, F. S., Gilmore, T. D. and Martindale, M. Q. (2013). "Microinjection of Mrna or Morpholinos for Reverse Genetic Analysis in the Starlet Sea Anemone, Nematostella Vectensis." Nature Protocols 8(5): 924-934.

Lee, E., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D. A., Jenkins, N. A. and Copeland, N. G. (2001). "A Highly Efficient *Escherichia Coli*-Based Chromosome Engineering System Adapted for Recombinogenic Targeting and Subcloning of Bac DNA." Genomics 73(1): 56-65.

Lee, P. Y., Nam, J. and Davidson, E. H. (2007). "Exclusive Developmental Functions of Gatae Cis-Regulatory Modules in the Strongylocentrorus Purpuratus Embryo." Developmental Biology 307(2): 434.

Love, A. C. and Raff, R. A. (2003). "Knowing Your Ancestors: Themes in the History of Evo,Äêdevo." Evolution & Development 5(4): 327-330.

Marlow, H., Roettinger, E., Boekhout, M. and Martindale, M. Q. (2012). "Functional Roles of Notch Signaling in the Cnidarian Nematostella Vectensis." Developmental biology 362(2): 295-308.

Martindale, M. Q. (2004). "Investigating the Origins of Triploblasty: `Mesodermal&Apos; Gene Expression in a Diploblastic Animal, the Sea Anemone Nematostella Vectensis (Phylum, Cnidaria; Class, Anthozoa)." Development (Cambridge, England) 131(10): 2463-2474.

Materna, S. C. and Davidson, E. H. (2012). "A Comprehensive Analysis of Delta Signaling in Pre-Gastrular Sea Urchin Embryos." Developmental biology 364(1): 77-87.

Matus, D. Q., Magie, C. R., Pang, K., Martindale, M. Q. and Thomsen, G. H. (2008). "The Hedgehog Gene Family of the Cnidarian, Nematostella Vectensis, and Implications for Understanding Metazoan Hedgehog Pathway Evolution." Developmental Biology 313(2): 501-518.

Matus, D. Q., Pang, K., Daly, M. and Martindale, M. Q. (2007a). "Expression of Pax Gene Family Members in the Anthozoan Cnidarian, Nematostella Vectensis." Evolution & Development 9(1): 25-38.

Matus, D. Q., Thomsen, G. H. and Martindale, M. Q. (2006). "Dorso/Ventral Genes Are Asymmetrically Expressed and Involved in Germ-Layer Demarcation During Cnidarian Gastrulation." Current Biology 16(5): 499-505.

Matus, D. Q., Thomsen, G. H. and Martindale, M. Q. (2007b). "Fgf Signaling in Gastrulation and Neural Development in Nematostella Vectensis, an Anthozoan Cnidarian." Development genes and evolution 217(2): 137-148.

McCauley, B. S., Weideman, E. P. and Hinman, V. F. (2010). "A Conserved Gene Regulatory Network Subcircuit Drives Different Developmental Fates in the Vegetal Pole of Highly Divergent Echinoderm Embryos." Developmental Biology 340(2): 200-208.

Medina, M. n., Collins, A. G., Silberman, J. D. and Sogin, M. L. (2001). "Evaluating Hypotheses of Basal Animal Phylogeny Using Complete Sequences of Large and Small Subunit Rrna." Proceedings of the National Academy of Sciences 98(17): 9707-9712.

Ohguro, Y., Takata, H. and Kominami, T. (2011). "Involvement of Delta and Nodal Signals in the Specification Process of Five Types of Secondary Mesenchyme Cells in Embryo of the Sea Urchin, Hemicentrotus Pulcherrimus." Development, Growth & Differentiation 53(1): 110-123.

Pankow, S. and Bamberger, C. (2007). "The P53 Tumor Suppressor-Like Protein Nvp63 Mediates Selective Germ Cell Death in the Sea Anemone Nematostella Vectensis." PLoS ONE 2(9): e782.

Passamaneck, Y. J. and Martindale, M. Q. (2012). "Cell Proliferation Is Necessary for the Regeneration of Oral Structures in the Anthozoan Cnidarian Nematostella Vectensis." BMC Developmental Biology 12(1): 1-13.

Prud'Homme, B., Gompel, N., Rokas, A., Kassner, V. A., Williams, T. M., Yeh, S.-D., True, J. R. and Carroll, S. B. (2006). "Repeated Morphological Evolution through Cis-Regulatory Changes in a Pleiotropic Gene." Nature 440(7087): 1050-1053.

Putnam, N., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E. and Kapitonov, V. (2007). "Sea Anemone Genome Reveals Ancestral Eumetazoan Gene Repertoire and Genomic Organization." Sci 317(5834): 86.

Reitzel, A. M., Behrendt, L. and Tarrant, A. M. (2010). "Light Entrained Rhythmic Gene Expression in the Sea Anemone Nematostella Vectensis: The Evolution of the Animal Circadian Clock." PLoS ONE 5(9): e12805.

Reitzel, A. M., Sullivan, J. C., Traylor-Knowles, N. and Finnerty, J. R. (2008). Genomic Survey of Candidate Stress-Response Genes in the Estuarine Anemone Nematostella Vectensis. The Biological Bulletin. 214**:** 233-254.

Renfer, E., Amon-Hassenzahl, A., Steinmetz, P. and Technau, U. (2010). "A Muscle-Specific Transgenic Reporter Line of the Sea Anemone, Nematostella Vectensis." PNAS 107(1): 104.

Rentzsch, F., Anton, R., Saina, M., Hammerschmidt, M., Holstein, T. W. and Technau, U. (2006). "Asymmetric Expression of the Bmp Antagonists Chordin and Gremlin in the Sea Anemone Nematostella Vectensis: Implications for the Evolution of Axial Patterning." Developmental Biology 296(2): 375-387.

Rentzsch, F., Fritzenwanker, J., Scholz, C. and Technau, U. (2008). "Fgf Signalling Controls Formation of the Apical Sensory Organ in the Cnidarian *Nematostella Vectensis*." Development 135(10): 1761.

Revilla-i-Domingo, R., Oliveri, P. and Davidson, E. H. (2007). "A Missing Link in the Sea Urchin Embryo Gene Regulatory Network: Hesc and the Double-Negative Specification of Micromeres." Proceedings of the National Academy of Sciences 104(30): 12383-12388.

Röttinger, E., Dahlin, P. and Martindale, M. Q. (2012). "A Framework for the Establishment of a Cnidarian Gene Regulatory Network for ‚Äúendomesoderm‚Äù Specification: The Inputs of √Ü-Catenin/Tcf Signaling." PLoS genetics 8(12): e1003164.

Rowan, S. and Cepko, C. L. (2004). "Genetic Analysis of the Homeodomain Transcription Factor Chx10 in the Retina Using a Novel Multifunctional Bac Transgenic Mouse Reporter." Developmental Biology 271(2): 388-402.

Ryan, J. F., Burton, P. M., Mazza, M. E., Kwong, G. K., Mullikin, J. C. and Finnerty, J. R. (2006). "The Cnidarian-Bilaterian Ancestor Possessed at Least 56 Homeoboxes: Evidence from the Starlet Sea Anemone, Nematostella Vectensis." Genome Biology 7(7): R64.

Saina, M., Genikhovich, G., Renfer, E. and Technau, U. (2009). "Bmps and Chordin Regulate Patterning of the Directive Axis in a Sea Anemone." Proc Natl Acad Sci US A 106(44): 18592-18597.

Scholz, C. B. and Technau, U. (2003). "The Ancestral Role of Brachyury: Expression of Nembra1 in the Basal Cnidarian Nematostella Vectensis (Anthozoa)." Development genes and evolution 212(12): 563-570.

Shaner, N. C., Steinbach, P. A. and Tsien, R. Y. (2005). "A Guide to Choosing Fluorescent Proteins." Nature methods 2(12): 905-909.

Shapiro, M. D., Bell, M. A. and Kingsley, D. M. (2006). "Parallel Genetic Origins of Pelvic Reduction in Vertebrates." Proceedings of the National Academy of Sciences 103(37): 13753-13758.

Smith, J. (2008). "A Protocol Describing the Principles of Cis-Regulatory Analysis in the Sea Urchin." Nat Protoc 3(4): 710-718.

Smith, J. and Davidson, E. (2008a). "Gene Regulatory Network Subcircuit Controlling a Dynamic Spatial Pattern of Signaling in the Sea Urchin Embryo." PNAS 105(51): 20089-20094.

Smith, J. and Davidson, E. (2009). "Regulative Recovery in the Sea Urchin Embryo and the Stabilizing Role of Fail-Safe Gene Network Wiring." PNAS 106(43): 18291-18296.

Smith, J. and Davidson, E. H. (2008b). "Gene Regulatory Network Subcircuit Controlling a Dynamic Spatial Pattern of Signaling in the Sea Urchin Embryo." Proceedings of the National Academy of Sciences of the United States of America 105(51): 20089-20094.

Smith, J. and Davidson, E. H. (2008c). "A New Method, Using *Cis*-Regulatory Control, for Blocking Embryonic Gene Expression." Developmental Biology 318(2): 360-365.

Sommer, R. J. (2009). "The Future of Evo,Äìdevo: Model Systems and Evolutionary Theory." Nature Reviews Genetics 10(6): 416-422.

Steele, R. E., David, C. N. and Technau, U. (2011). "A Genomic View of 500 Million Years of Cnidarian Evolution." Trends in genetics : TIG 27(1): 7-13.

Stefanik, D. J., Friedman, L. E. and Finnerty, J. R. (2013a). "Collecting, Rearing, Spawning and Inducing Regeneration of the Starlet Sea Anemone, Nematostella Vectensis." Nature Protocols 8(5): 916-923.

Stefanik, D. J., Wolenski, F. S., Friedman, L. E., Gilmore, T. D. and Finnerty, J. R. (2013b). "Isolation of DNA, Rna and Protein from the Starlet Sea Anemone Nematostella Vectensis." Nature Protocols 8(5): 892-899.

Stern, D. L. (1998). "A Role of Ultrabithorax in Morphological Differences between Drosophila Species." Nature 396(6710): 463.

Tessmar-Raible, K., Steinmetz, P. R. H., Snyman, H., Hassel, M. and Arendt, D. (2005). "Fluorescent Two-Color Whole Mount in Situ Hybridization in *Platynereis Dumerilii* (Polychaeta, Annelida), an Emerging Marine Molecular Model for Evolution and Development." Biotechniques 39: 460-462.

Tulin, S., Aguiar, D., Istrail, S. and Smith, J. (2013). "A Quantitative Reference Transcriptome for Nematostella Vectensis Early Embryonic Development: A Pipeline for De Novo Assembly in Emerging Model Systems." EvoDevo 4(1): 16.

Wahl, M. E., Hahn, J., Gora, K., Davidson, E. H. and Oliveri, P. (2009). "The Cis-Regulatory System of the *Tbrain* Gene: Alternative Use of Multiple Modules to Promote Skeletogenic Expression in the Sea Urchin Embryo." Developmental Biology 335(2): 428-441.

Wikramanayake, A., Hong, M., Lee, P., Pang, K., Byrum, C., Bince, J., Xu, R. and Martindale, M. (2003). "An Ancient Role for Nuclear -Catenin in the Evolution of Axial Polarity and Germ Layer Segregation." Nature 426(6965): 446-450.

Wolenski, F. S., Layden, M. J., Martindale, M. Q., Gilmore, T. D. and Finnerty, J. R. (2013). "Characterizing the Spatiotemporal Expression of Rnas and Proteins in the Starlet Sea Anemone, Nematostella Vectensis." Nature Protocols 8(5): 900-915.

Wray, G. A. (2007). "The Evolutionary Significance of Cis-Regulatory Mutations." Nature Reviews Genetics 8(3): 206-216.

**Table 1. Sequenced BACs identified through a BAC filter screen, and their ID, genomic location, and total length inferred from the genome.**

| Identified through BAC filter screen: | Scaffold(s) | BAC ID | Total length of the BAC inferred from the distance between BAC-end sequences within the current genome assembly in bp |
|---|---|---|---|
| *Nv-chordin* | 71; 677 | 12-k-13 | 172637 |
| *Nv-foxa* | 58; 2433 | 8-k-22 | 178878 |
| *Nv-dpp* | 106; 4457 | 2-g-22 | 149307 |
| *Nv-wnta* | 28; 46; +8 more | 14-g-24 | 113509 |
| *Nv-wnt1* | 3; 8; +7 more | 12-H-10 | 128531 |

**Table 2. BACs identified through BLASTing BAC-end sequences against the Nematostella genome, their ID, total length as inferred from the genome and their genomic location.**

| Identified through BAC-end alignment: | Scaffold | BAC ID | Total length of the the BAC inferred from the distance between BAC-end sequences within the current genome assembly in bp | Beginning at position | Ending at position |
|---|---|---|---|---|---|
| *Nv-brachyury* | 137 | 24d6 | 156049 | 137512 | 293561 |
| *Nv-gremlin* | 39 | | 151089 | 1055085 | 1206174 |
| *Nv-snailB* | 5 | 38n15 | 185773 | 2054413 | 2240186 |
| *Nv-tolloid* | 67 | 3d13 | 166942 | 820992 | 654050 |

**Table 3. Number of GFP expressing larva after injecting *Nv-bra*:eGFP::L10, *Nv-dpp*:eGFP::L10 and *Nv-grm*:eGFP::L10 in three independent experiments for each BAC, showing the developmental stage, total number of analyzed larvae, the total number of GFP-expressing larvae and that number in percent, the number of larvae showing faithful mosaic expression, and that number in percent as well as the number of larvae with ectopic expression and that number in percent.**

| Experiment | Injected BAC | Total number of larvae analyzed | total number of larvae expressing GFP after 2 days | *No. of larvae expressing GFP after 2 days in %* | total number of larvae with faithful mosaic expression | *faithful mosaic expression in %* | total number of larvae with ectopic expresion | *ectopic expression in %* |
|---|---|---|---|---|---|---|---|---|
| **1** | *Nv-Bra:eGFP::L10 BAC* | 112 | 15 | *13.4* | 7 | *46.7* | 8 | *53.3* |
| **2** | | 170 | 15 | *8.8* | 6 | *40.0* | 9 | *60.0* |
| **3** | | 102 | 13 | *12.7* | 5 | *38.5* | 8 | *61.5* |
| **total** | | **384** | **43** | ***11.2*** | **18** | ***41.9*** | **25** | ***58.1*** |
| **1** | *Nv-dpp*:eGFP::L10 BAC | 75 | 6 | *8.0* | 3 | *50.0* | 3 | *50.0* |
| **2** | | 76 | 8 | *10.5* | 1 | *12.5* | 7 | *87.5* |
| **3** | | 156 | 18 | *11.5* | 6 | *33.3* | 12 | *66.7* |
| **total** | | **307** | **32** | ***10.4*** | **10** | ***31.3*** | **22** | ***68.8*** |
| **1** | *Nv-grm*:eGFP::L10 BAC | 86 | 10 | *11.6* | 2 | *20.0* | 8 | *80.0* |
| **2** | | 183 | 27 | *14.8* | 7 | *25.9* | 20 | *74.1* |
| **3** | | 108 | 28 | *25.9* | 8 | *28.6* | 20 | *71.4* |
| **total** | | **377** | **65** | ***17.2*** | **17** | ***26.2*** | **48** | ***73.8*** |

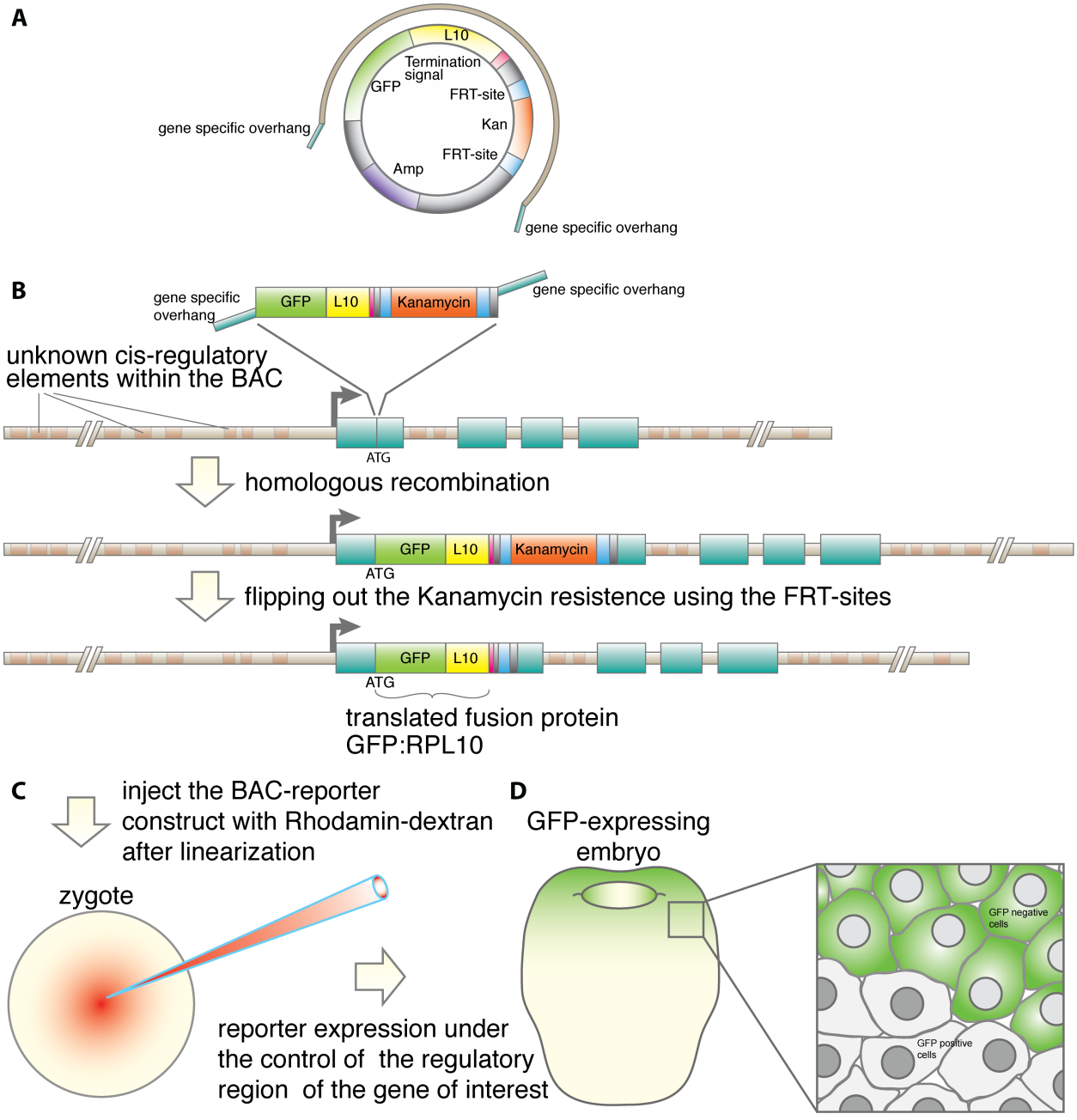


**Figure 1. BAC reporter construction via homologous recombination.**
(A) Vector containing the recombination cassette composed of eGFP::*Nv-L10,* 374 bp 3'UTR of L10 including the termination signal, the Kanamycin resistance gene *KmR* flanked by FRT-sites and Ampicilin resistance gene *AmpR.* To generate the recombination cassette, we used PCR primers with 45bp overhangs specific to the insertion site being targeted in the gene ("gene-specific overhang"). (B) The amplified recombination cassette is inserted in frame at the start codon of the GOI within the BAC by homologous recombination. After successful integration the Kanamycin resistance is flipped out resulting in a reporter construct with the GFP::L10 fusion protein expressed under the identical regulatory control of the GOI. (C) The recombineered BAC is linearized and injected into zygotes together with an injection tracer such as Rhodamine

dextran. (D) The larva expresses the GFP-L10 fusion protein in cells where the GOI would be expressed.

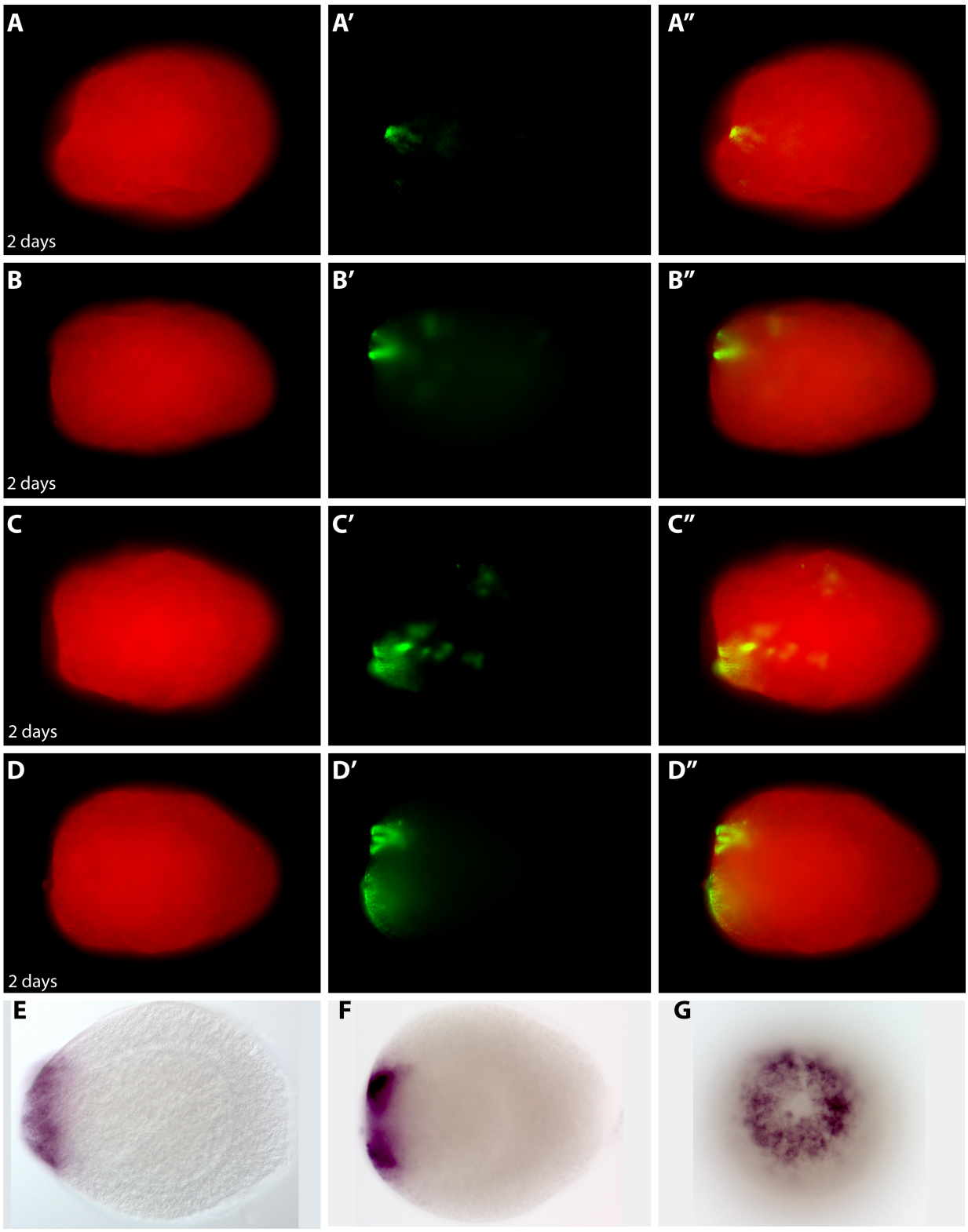

**Figure 2. Examples of *Nv-bra:GFP::L10* expression and WMISH against *Nv-bra*.**
Each row represents a single larva: left column shows red channel only (Rhodamine dextran, our injection tracer), middle column, green channel (eGFP) and right column, overlay between the red and the green channel. Live planula larvae were imaged at 2 days post fertilization, blastopore oriented to the left, except F. (A-B'') Larvae with one eGFP expression domain near the blastopore. (C-C'') A planula shows one eGFP expression domain in the blastopore region and additional eGFP positive cells in the middle of the ectoderm. (D-D'') A planula shows two eGFP expressing domains near the blastopore, possibly the result of two independent integration events. D-E: WMISH against *Nv-bra* in 2 days old planula, F view on the blastopore. A detailed description of the expression pattern is provided in the results section.

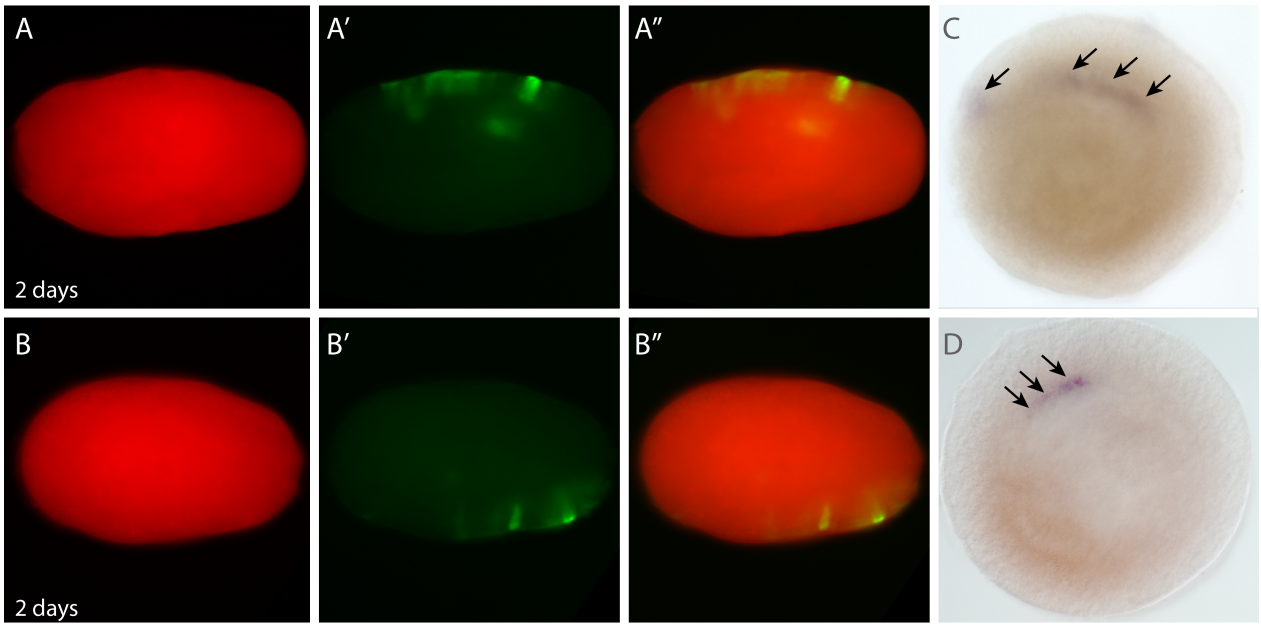


**Figure 3: Examples of *Nv-dpp:GFP::L10* expression and WMISH against *Nv-dpp*.**
Left column shows red channel only (Rhodamine dextran, our injection tracer), middle column, green channel (eGFP) and the second to the right column shows the overlay between the red and the green channel of the same larave. Live planula larvae were imaged at 2 days post fertilization. The blastopore is oriented to the left in all images except D. (A-A'') Planula shows GFP expression on one side in several small domains in a stipe-like area along the oral-aboral axis. While most of the GFP positive cells seem to be ectodermal (columnar epithelial cells), activity in endodermal cells can be seen directly underneath. (B-B'') Planula has GFP expressing cells arranged in a stripe-like fashion along the oral-aboral axis. Most GFP positive cells seem to be located in the ectoderm. C and D: WMISH against *Nv-dpp* in 2 days old planula, D view on the blastopore. The arrows indicate the WMISH staining. A detailed description of the expression pattern is provided in the results section.

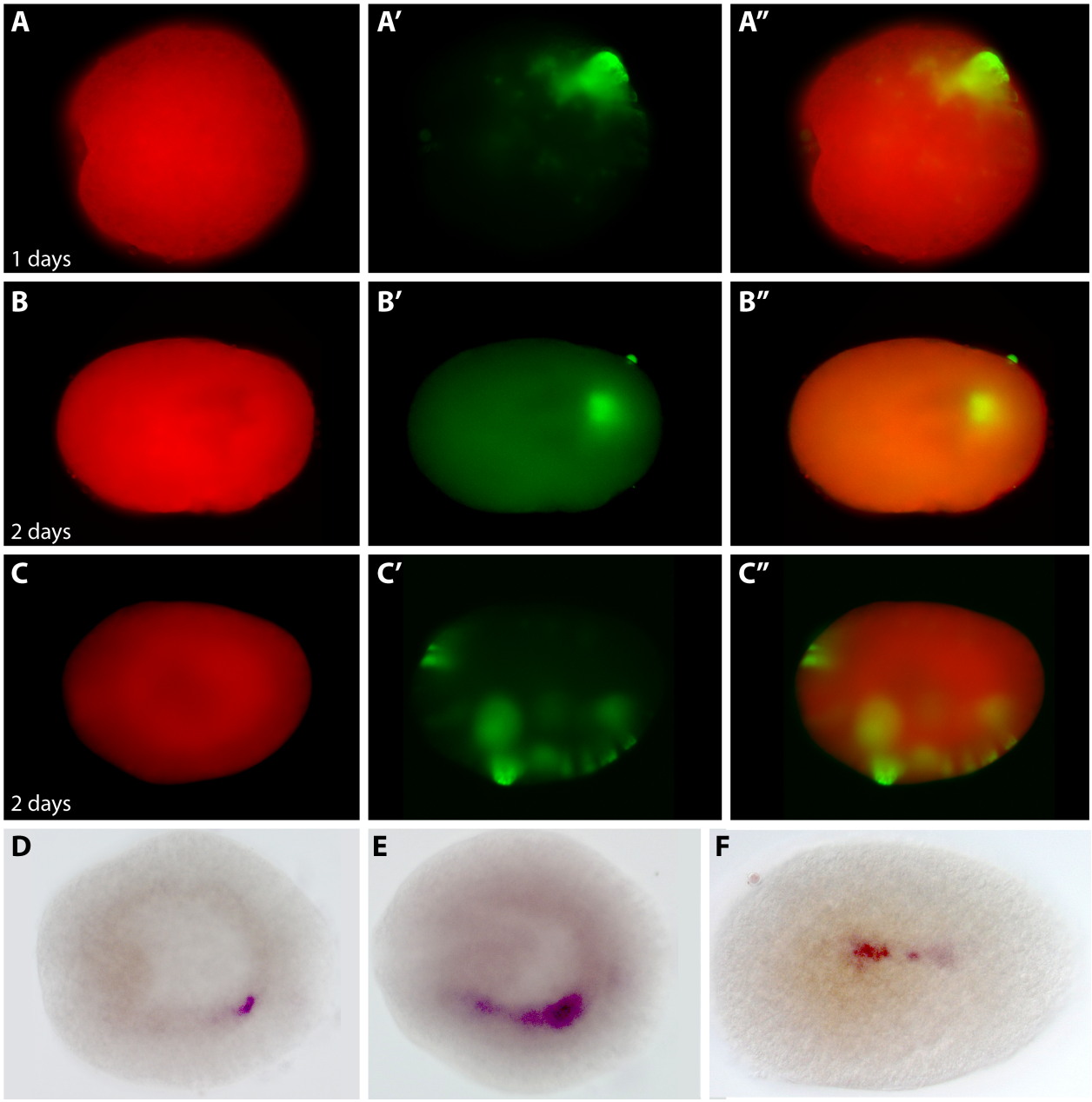


**Figure 4: Examples of *Nv-grm:GFP::L10* expression and WMISH against *Nv-grm*.**
Each row represents a single larva: left column shows red channel only (Rhodamine dextran, our injection tracer), middle column, green channel (eGFP) and right column, overlay between the red and the green channel. Planula larvae were imaged alive at either 1 day post fertilization (A-A'') or 2 days post fertilization (B-C''), blastopore to the left. (A-A'') An injected planula larva shows GFP expressing domain is located on one side, near the aboral end of the larva, including ectodermal and endodermal cells. (B-B'') A planula larva shows GFP expression in an endodermal-only, domain on one side near the aboral end. (C-C'') Planula larva shows GFP expression in several domains: two small domains near the blastopore; several small ectodermal domains in a stripe along the oral-aboral axis and at least 3 endodermal domains in that same strip-like domain. D-E: WMISH against *Nv-grm* in 2 days old planula. The arrows indicate the WMISH staining. A detailed description of the expression pattern is provided in the results

section.